

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2003, 18:27:17 ; Search time 28.9091 Seconds
(without alignments)
107.116 Million cell updates/sec

Title: us-09-743-225-8
Perfect score: 58
Sequence: 1 NTLKTPRVGGXA 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL-23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 40 | 69.0 | 268 | 5 Q9XZ5 | Q9XZ5 leishmania |
| 2 | 38 | 65.5 | 334 | 2 Q9F937 | Q9F937 pseudomonas |
| 3 | 38 | 65.5 | 742 | 3 Q9HG15 | Q9HG15 yarrowia 11 |
| 4 | 38 | 65.5 | 991 | 3 Q12252 | Q12252 saccharomyc |
| 5 | 37 | 63.8 | 388 | 16 Q8G3N7 | Q8G3N7 bifidobacte |
| 6 | 36 | 62.1 | 257 | 16 Q9CGF3 | Q9CGF3 lactococcus |
| 7 | 36 | 62.1 | 266 | 10 Q9FZ93 | Q9FZ93 arabisdopsi |
| 8 | 36 | 62.1 | 265 | 10 Q8GWE8 | Q8GWE8 arabisdopsi |
| 9 | 36 | 62.1 | 319 | 10 Q9LRT3 | Q9LRT3 shewanella |
| 10 | 36 | 62.1 | 409 | 16 Q8EIH7 | Q8EIH7 haemophilus |
| 11 | 36 | 62.1 | 1002 | 2 Q8GM78 | Q8GM78 haemophilus |
| 12 | 36 | 62.1 | 1004 | 2 Q8GM77 | Q8GM77 haemophilus |
| 13 | 36 | 62.1 | 1301 | 12 Q9EMI5 | Q9EMI5 ansacta moo |
| 14 | 35 | 60.3 | 132 | 4 Q8N8V9 | Q8N8V9 homo sapien |
| 15 | 35 | 60.3 | 143 | 12 Q8BC65 | Q8BC65 human papil |
| 16 | 35 | 60.3 | 149 | 16 Q98M54 | Q98M54 rhizobium 1 |

| | | | | | |
|----|----|------|------|-----------|--------------------|
| 17 | 35 | 60.3 | 235 | 16 Q8ZRN2 | Q8ZRN2 salmonella |
| 18 | 35 | 60.3 | 235 | 16 Q8Z993 | Q8Z993 salmonella |
| 19 | 35 | 60.3 | 235 | 16 Q8X8W1 | Q8X8W1 escherichia |
| 20 | 35 | 60.3 | 306 | 16 Q8FM02 | Q8FM02 corynebacte |
| 21 | 35 | 60.3 | 310 | 5 Q20160 | Q20160 caenorhabdi |
| 22 | 35 | 60.3 | 324 | 16 Q8P8P9 | Q8P8P9 xanthomonas |
| 23 | 35 | 60.3 | 339 | 2 Q8KI09 | Q8KI09 pseudomonas |
| 24 | 35 | 60.3 | 339 | 2 Q8KIT9 | Q8KIT9 pseudomonas |
| 25 | 35 | 60.3 | 340 | 16 Q8PEV4 | Q8PEV4 xanthomonas |
| 26 | 35 | 60.3 | 341 | 2 Q8GHQ4 | Q8GHQ4 pseudomonas |
| 27 | 35 | 60.3 | 353 | 16 Q9KDZ1 | Q9KDZ1 bacillus ha |
| 28 | 35 | 60.3 | 357 | 16 Q66804 | Q66804 aquifex ase |
| 29 | 35 | 60.3 | 413 | 16 Q8ZHH9 | Q8ZHH9 yersinia pe |
| 30 | 35 | 60.3 | 447 | 16 Q8CZW0 | Q8CZW0 yersinia pe |
| 31 | 35 | 60.3 | 456 | 10 Q8L8A1 | Q8L8A1 arabidopsis |
| 32 | 35 | 60.3 | 456 | 10 Q8LEU4 | Q8LEU4 arabidopsis |
| 33 | 35 | 60.3 | 456 | 10 Q8GWW8 | Q8GWW8 arabidopsis |
| 34 | 35 | 60.3 | 461 | 10 Q9FI35 | Q9FI35 arabidopsis |
| 35 | 35 | 60.3 | 514 | 5 Q21458 | Q21458 caenorhabdi |
| 36 | 35 | 60.3 | 530 | 10 Q8H979 | Q8H979 rhus vernic |
| 37 | 35 | 60.3 | 533 | 10 Q94ID0 | Q94ID0 rhus vernic |
| 38 | 35 | 60.3 | 550 | 16 Q8XR68 | Q8XR68 raietonia s |
| 39 | 35 | 60.3 | 637 | 17 Q972X6 | Q972X6 sulfobolus |
| 40 | 35 | 60.3 | 1592 | 5 Q01583 | Q01583 caenorhabdi |
| 41 | 35 | 60.3 | 2591 | 2 Q54959 | Q54959 streptomyce |
| 42 | 34 | 58.6 | 127 | 3 Q9HEU8 | Q9HEU8 emericeila |
| 43 | 34 | 58.6 | 177 | 4 Q8N7V0 | Q8N7V0 homo sapien |
| 44 | 34 | 58.6 | 185 | 5 Q9NKU2 | Q9NKU2 leishmania |
| 45 | 34 | 58.6 | 229 | 5 Q8MPN9 | Q8MPN9 drosophila |

ALIGNMENTS

RESULT 1

Q9XZ5 PRELIMINARY; PRT; 268 AA.
AC Q9XZ5
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical 30.3 kDa protein.
GN L2743.07.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL389894; CAC22688.1;
RW Hypothetical protein.
SQ SEQUENCE 268 AA; 30293 MW; D68B62C0ADD5444 CRC64;

Query Match 69.0%; Score 40; DB 5; Length 268;

Best Local Similarity 66.7%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NTLKTPRVGGXA 12

Db 200 NPKTPRTGGAA 211

RESULT 2

Q9F937 PRELIMINARY; PRT; 334 AA.
 ID Q9F937
 AC Q9F937
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Glycosyltransferase L.
 GN WPPL
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KT2440;
 RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.L.;
 RT "Characterization of a Pseudomonas putida wbpL mutant."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF209871; AAG12993.1;
 DR InterPro; IPR000715; Glyco.trans.4.
 DR Pfam; PF00953; Glycos_transf_4.1.
 KW transferase.
 SQ SEQUENCE 334 AA; 35637 MW; 1B9372E230A9F842 CRC64;

Query Match 65.5%; Score 38; DB 2; Length 334;
 Best Local Similarity 66.7%; Pred. No. 34;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTLKTPRVGGXA 12
 DB 38 HTLPTPRGGGLA 49
 :||| ||| |||

RESULT 3

Q9HG15 PRELIMINARY; PRT; 742 AA.
 ID Q9HG15
 AC Q9HG15
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Polypeptide release factor 3.
 GN SUP35.
 OS Yarrowia lipolytica (Candida lipolytica).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=4952;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21324710; PubMed=11430816;
 RA Nakayashiki T., Ebihara K., Bannai H., Nakamura Y.;
 RT "Yeast [PSI+] 'prions' that are cross-transmissible and susceptible
 RT beyond a species barrier through a quasi-prion state."
 RL Mol. Cell 7:1121-1130(2001).
 DR EMBL; AB039752; BAB12683.1;
 DR InterPro; IPR004160; EFTU_Cterm.
 DR InterPro; IPR004161; EFTU_D2.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR PRINTS; PR00315; ELONGATNFCT.
 KW GTP-binding; Protein biosynthesis.
 SQ SEQUENCE 742 AA; 80125 MW; 3C8D3F5738206C0A CRC64;

Query Match 65.5%; Score 38; DB 3; Length 742;
 Best Local Similarity 70.0%; Pred. No. 79;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTLKTPRVGG 10
 DB 172 NKLKPRVGG 181
 | | | | |

RESULT 4

Q12252 PRELIMINARY; PRT; 991 AA.
 ID Q12252
 AC Q12252
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE ORF YOL084W from chromosome XV.
 GN PHM7 OR YOL084W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FY1679;
 RX MEDLINE=96021609; PubMed=8533473;
 RA Zumbstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
 RT "A 29.425 kb segment on the left arm of yeast chromosome XV contains
 RT more than twice as many unknown as known open reading frames."
 RL Yeast 11:975-986(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zumbstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X83121; CAA58195.1;
 DR EMBL; Z74826; CAA99096.1;
 DR SGD; S0005444; PHM7.
 DR InterPro; IPR003864; DUF221.
 DR Pfam; PF02714; DUF221; 1.
 SQ SEQUENCE 991 AA; 112545 MW; 491FAB0BD143DC5E CRC64;

Query Match 65.5%; Score 38; DB 3; Length 991;
 Best Local Similarity 77.8%; Pred. No. 11e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTLKTPRVG 9
 DB 569 NTLATPRMG 577
 ||| ||| |

RESULT 5

Q8G3N7 PRELIMINARY; PRT; 388 AA.
 ID Q8G3N7
 AC Q8G3N7
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Possible undecaprenyl-phosphate
 DE alpha-N-acetylglucosaminyltransferase.
 GN RFE OR BL1721
 OS Bifidobacterium longum.
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 OC Bifidobacteriaceae; Bifidobacterium.
 OX NCBI_TaxID=216816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCC 2705;
 RA MEDLINE=22294977; PubMed=12381787;
 RA Scheil M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
 RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
 RA Pridmore R.D., Arigoni F.;
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation
 RT to the human gastrointestinal tract."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
 DR EMBL; AE014806; AAN25505.1;
 KW Complete proteome.
 SQ SEQUENCE 388 AA; 41434 MW; COA14679F63A1DD5 CRC64;

Query Match 63.8%; Score 37; DB 16; Length 388;
Best Local Similarity 60.08; Pred. No. 63;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTLKTPRVGG 10
: : : : :
Db 2 HTIPTPRMG 11

RESULT 6

Q9CGF3 Q9CGF3 PRELIMINARY; PRT; 257 AA.
AC Q9CGF3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE N-acetylglucosamine catabolic protein.
GN YLPH OR L1143.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
[1]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL: AE006346; AAK05241.1; -
DR InterPro: IPR006357; HAD_SF_IIA.
DR InterPro: IPR006354; HAD_SF_IIA_hypl.
DR Pfam: PF00702; Hydrolase.
DR TIGRFAMs: TIGR01460; HAD_SF-IIA; 1.
DR TIGRFAMs: TIGR01457; HAD_SF-IIA-hyp2; 1.
KW Complete proteome.
SQ SEQUENCE 257 AA; 28179 MW; 75EC096A3C19BC3 CRC64;

Query Match 62.1%; Score 36; DB 16; Length 257;
Best Local Similarity 87.5%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTLKTPRV 8
: : : : :
Db 47 NTLKTPRV 54

RESULT 7

Q9FZ93 Q9FZ93 PRELIMINARY; PRT; 266 AA.
AC Q9FZ93
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE F3H9.15 protein.
GN F3H9.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN RN
RP SEQUENCE FROM N.A.
RA Federapfel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chlou J., Choi E., Gonzalez A.,
RA Hong B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharly N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vayaberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (AUG-2000) to the ENBL/GenBank/DBJ databases.

DR EMBL: AC021044; AAF98436.1; -
SQ SEQUENCE 266 AA; 30336 MW; 79F0D0AA1374D4BD CRC64;

Query Match 62.1%; Score 36; DB 10; Length 266;
Best Local Similarity 77.8%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLKTPRVGG 10
: : : : :
Db 115 TLKTPQNG 123

RESULT 8

Q8GWE8 Q8GWE8 PRELIMINARY; PRT; 266 AA.
AC Q8GWE8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN AtIG28190/F3H9.13
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Arabidopsis thaliana full-length cDNA.";
RL Submitted (NOV-2002) to the ENBL/GenBank/DBJ databases.
DR EMBL: AK118888; BAC43473.1; -
KW Hypothetical protein.
SQ SEQUENCE 266 AA; 30304 MW; 39F0DCBC8196266 CRC64;

Query Match 62.1%; Score 36; DB 10; Length 266;
Best Local Similarity 77.8%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLKTPRVGG 10
: : : : :
Db 115 TLKTPQNG 123

RESULT 9

Q9LTT3 Q9LTT3 PRELIMINARY; PRT; 319 AA.
AC Q9LTT3
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE High mobility group protein-like (Atg13350/MDC11_14) (Hypothetical
DE protein).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the ENBL/GenBank/DBJ databases.
[2]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty p1 and TAC

```

RT  clones."
RL  DNA Res. 7:131-135(2000).
RN  [3]
RP  SEQUENCE FROM N.A.
RA  Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
RA  Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
RA  Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
RA  Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA  Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
RA  Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W.,
RA  Theologis A., Ecker J.R.;
RT  "Arabidopsis cDNA clones.";
RL  Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN  [4]
RP  SEQUENCE FROM N.A.
RA  Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA  Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT  "Full-length messenger RNA sequences greatly improve genome
RT  annotation.";
RL  Genome Biol. 0:0-0(2002).
RN  [5]
RP  SEQUENCE FROM N.A.
RA  Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA  Feldmann K.;
RT  "Full-length cDNA from Arabidopsis thaliana.";
RL  Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN  [6]
RP  SEQUENCE FROM N.A.
RA  Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
RA  Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA  Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA  Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA  Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA  Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA  Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT  "Arabidopsis ORF clones.";
RL  Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN  [7]
RP  SEQUENCE FROM N.A.
RA  EMBL; AB024034; BAB02804.1; -
DR  EMBL; AF361582; AAK32750.1; -
DR  EMBL; AY087770; AAM65305.1; -
DR  EMBL; AY133557; AAM91387.1; -
DR  HSP; P07135; IHMF.
DR  InterPro; IPR001606; ARID.
DR  InterPro; IPR000910; HMG_12_box.
DR  Pfam; PF01388; ARID; 1.
DR  Pfam; PF00505; HMG_box; 1.
KW  Hypothetical protein.
SQ  SEQUENCE 319 AA; 36296 MW; 7802A12A69F7456C CRC64;

Query Match 62.1%; Score 36; DB 10; Length 319;
Best Local Similarity 77.8%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLKTPRVGG 10
Db 60 TLKVPVGG 68

RESULT 10
Q8EIH7 PRELIMINARY; PRT; 409 AA.
ID Q8EIH7
AC Q8EIH7
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase.
GN SERA OR S00862.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.

clones."
DNA Res. 7:131-135(2000).
[3]
SEQUENCE FROM N.A.
Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W.,
Theologis A., Ecker J.R.;
"Arabidopsis cDNA clones.";
Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
"Full-length messenger RNA sequences greatly improve genome
annotation.";
Genome Biol. 0:0-0(2002).
[5]
SEQUENCE FROM N.A.
Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
Feldmann K.;
"Full-length cDNA from Arabidopsis thaliana.";
Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[6]
SEQUENCE FROM N.A.
Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
"Arabidopsis ORF clones.";
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE FROM N.A.
EMBL; AB024034; BAB02804.1; -
DR EMBL; AF361582; AAK32750.1; -
DR EMBL; AY087770; AAM65305.1; -
DR EMBL; AY133557; AAM91387.1; -
DR HSP; P07135; IHMF.
DR InterPro; IPR001606; ARID.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF01388; ARID; 1.
DR Pfam; PF00505; HMG_box; 1.
KW Hypothetical protein.
SQ SEQUENCE 319 AA; 36296 MW; 7802A12A69F7456C CRC64;

Query Match 62.1%; Score 36; DB 10; Length 319;
Best Local Similarity 77.8%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLKTPRVGG 10
Db 60 TLKVPVGG 68

RESULT 10
Q8EIH7 PRELIMINARY; PRT; 409 AA.
ID Q8EIH7
AC Q8EIH7
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase.
GN SERA OR S00862.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-MR-1;
RX MEDLINE-22297686; PubMed-12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015531; AAN53938.1; -
DR TIGR; S00862; -
KW Complete proteome.
SQ SEQUENCE 409 AA; 44314 MW; DCA4DD5D3AD7E2B0 CRC64;

Query Match 62.1%; Score 36; DB 16; Length 409;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NTLKTPRVGG 10
Db 286 NVLLTPHVG 295

RESULT 11
Q8GM78 PRELIMINARY; PRT; 1002 AA.
ID Q8GM78
AC Q8GM78
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Adhesin.
GN HIA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-nontypable strain 3248A;
RX MEDLINE-22300350; PubMed-12410830;
RA Laarmann S., Cutter D., Juehne T., Barenkamp S.J., St Geme J.W.;
RT "The Haemophilus influenzae Hia autotransporter harbours two adhesive
RT pockets that reside in the passenger domain and recognize the same
RT host cell receptor.";
RL Mol. Microbiol. 46:731-743(2002).
DR EMBL; AY078085; AAL79951.1; -
SQ SEQUENCE 1002 AA; 103638 MW; F9C4130DC3C37EAE CRC64;

Query Match 62.1%; Score 36; DB 2; Length 1002;
Best Local Similarity 70.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NTLKTPRVGG 10
Db 222 STLDDPRVGG 231

RESULT 12
Q8GM77 PRELIMINARY; PRT; 1004 AA.
ID Q8GM77
AC Q8GM77
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Adhesin.
GN HIA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.

```

OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=non-typeable strain 1862A;
 RX MEDLINE=22300350; PubMed=12410830;
 RA Laarmann S., Cutter D., Juehne T., Barenkamp S.J., St Geme J.W.;
 RT "The Haemophilus influenzae Hia autotransporter harbours two adhesive
 RT pockets that reside in the passenger domain and recognize the same
 RT host cell receptor";
 RL Mol. Microbiol. 46:731-743(2002).
 DR EMBL: AY078086; AAL79521.1;
 SQ SEQUENCE 1004 AA; 104286 MW; 78067E2D4E3D017 CRC64;

 Query Match 62.1%; Score 36; DB 2; Length 1004;
 Best Local Similarity 70.0%; Pred. No. 2.7e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 Qy 1 NTLKTPRVGG 10
 Db 224 STLDPRVGG 233

 RESULT 13
 Q9EMI5
 ID Q9EMI5 PRELIMINARY; PRT; 1301 AA.
 AC Q9EMI5;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE AMV221.
 GN AMV221.
 OS Anaseta moorei entomopoxvirus (AnEPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
 OC Entomopoxvirus B.
 OX NCBI_TaxID=28321;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20396580; PubMed=10936094;
 RA Bowden A.L., Glassberg K.J., Diggins J., Shaw R., Farmerie W.,
 RA Moyer R.W.;
 RT "Complete Genomic Sequence of the Anaseta moorei Entomopoxvirus:
 RT Analysis and Comparison with Other Poxviruses";
 RL Virology 274:120-139(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bowden A.L., Glassberg K.J., Diggins J., Shaw R., Farmerie W.,
 RA Moyer R.W.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF250284; AAG02927.1;
 DR InterPro: IPR006592; RNA_pol_A.
 DR InterPro: IPR000722; RNA_pol_A.
 DR InterPro: IPR007066; RNA_pol_Rpb1_3.
 DR InterPro: IPR007083; RNA_pol_Rpb1_4.
 DR Pfam: PF00623; RNA_pol_Rpb1_2; 1.
 DR Pfam: PF04983; RNA_pol_Rpb1_3; 1.
 DR Pfam: PF05000; RNA_pol_Rpb1_4; 1.
 DR SMART: SM00663; RPOA_N; 1.
 SQ SEQUENCE 1301 AA; 150298 MW; 3E7D097185975B7 CRC64;

 Query Match 62.1%; Score 36; DB 12; Length 1301;
 Best Local Similarity 50.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

 Qy 1 NTLKTPRVGG 10
 Db 30 NSVSPKLG 39

 RESULT 14
 Q8NBV9
 ID Q8NBV9 PRELIMINARY; PRT; 132 AA.
 AC Q8NBV9;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)

DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ38783.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahara K.,
 RA Masuho Y., Nagai K., Isogai T.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK096102; BAC04702.1;
 KW Hypothetical protein.
 SQ SEQUENCE 132 AA; 14581 MW; 22DA60CC890A8211 CRC64;

 Query Match 60.3%; Score 35; DB 4; Length 132;
 Best Local Similarity 60.0%; Pred. No. 51;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

 Qy 1 NTLKTPRVGG 10
 Db 65 STLASPLRG 74

 RESULT 15
 Q8BC65
 ID Q8BC65 PRELIMINARY; PRT; 143 AA.
 AC Q8BC65;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Major capsid protein (fragment).
 GN Li.
 OS Human papillomavirus.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10566;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FA93;
 RA Antonsson A., Holmgren V., Hansson B.G.;
 RT "Prevalence and type spectrum of human papillomavirus in normal skin
 RT from three continents";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF542099; AAN28674.1;
 FT NON_TER 1
 FT NON_TER 143
 SQ SEQUENCE 143 AA; 16103 MW; 7B2EF7C742948DBD CRC64;

 Query Match 60.3%; Score 35; DB 12; Length 143;
 Best Local Similarity 60.0%; Pred. No. 56;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 Qy 1 NTLKTPRVGG 10
 Db 8 NTLEVPKVG 17

Search completed: August 28, 2003, 18:37:59
 Job time : 30.9091 secs